Attachment E

USDA-ARS

Information Technology Management
An Information/Communication Strategy for National Programs

July 17, 1998
Final Report

By the

ARS Information Technology Management
National Programs Information Strategy Action Team

Claudia Weston, Chair,
and
Ray Carruthers, Sandy Hays, Bruce Kinzel, Pamela Mason, Lee Panella, Gail Poulos, and Dave Rust

Executive Summary

Information Technology Action Team # 5 (ITM 5) was charged with developing an overall strategy for communicating the National Programs to customers, stakeholders and ARS staff and integrating the related program elements utilizing electronic technologies. The resulting strategy would then be used to implement an integrated communication process whereby information related to ARS research activities would be made more accessible. This team developed a number of short term action items which it felt could and should be accomplished within a limited time frame. These action items included: establishing a National Program Staff web presence; soliciting feedback and comments on the National Program descriptions by alerting ARS' customers and stakeholders to the presence of these descriptions on the ARS web site; and, once the National Program descriptions were finalized, creating a brochure aimed at ARS staff outlining the new structure of the National Program Staff. Developing an integrated communication process whereby information related to ARS research activities would be made more accessible was identified as the sole long term action item.

Proposed activities related to this communication process were limited in scope to the creation of Internet-accessible resources made available through the ARS web site. These recommended resources focus on research activities from an ARS-wide perspective thereby enabling individual areas and locations to concentrate on the topics and concerns of their unique customer base. These resources include: predefined descriptions of ARS activities in terms of specific commodities (such as corn, potatoes, dairy products, etc.) or technologies (such as genetic engineering, pesticide reduction, etc.); matrices linking aspects of the National Programs to the actual research locations; a dynamic contact list of experts; responses to frequently asked questions (FAQs) and a section on hot topics; and a centralized access mechanism for Internet-accessible ARS-generated research publications, data, and information. The dynamic contact list of experts has already been developed and is available via the "Find the Expert" database located on the ARS web site.

Contents

Background 2

Vision and Mission of the Agricultural Research Service 3

ARS Headquarters 3

Centralized Access Mechanism 4
- The Centralized System 6
- The Hybrid System 6
- The Decentralized System 6
- System Recommendation 7

Summary of Action Item Recommendations and Activities 7

Appendix A
- Schedule & Milestones 8

Appendix B
- Background Material 9

Note: Appendices available upon request.

Background

During the September Administrative Council (AC) meeting, the topic of information technology management in USDA and ARS was discussed and a number of action teams formed. Of these action teams, the fifth one was charged with developing an overall strategy for communicating the National Programs to customers, stakeholders and ARS staff and integrating the related program elements utilizing electronic technologies. This charge was modified slightly during the December AC meeting to limit the scope to include external customers only. The resulting strategy would then be used to implement an integrated communication process whereby information related to ARS research activities would be made more accessible. The original action team consisted of staff of the National Program Staff (NPS), Information Staff (IS), Office of Technology Transfer (OTT), Administrative and Financial Management (AFM), a field location (the Northern Plains Area (NPA)), and the National Agricultural Library (NAL) and included: Pamela Mason (NPS), Dave Rust (NPS), Ray Carruthers (NPS), Sandy Miller Hays (IS), Bruce Kinzel (OTT), Gail Poulos (OTT), Dave Carter (AFM), Lee Panella (NPA), and Claudia Weston (NAL) as chair.

A two pronged approach was taken to meet both the short term and long term external communication needs of the National Program Staff (NPS). To fulfill the short term needs, staff from NPS, IS, and NAL worked together to establish an NPS web presence. This web presence included descriptions of the National Programs and the "Find the Expert" database.

An announcement was sent to the Federal Register soliciting feedback on the new National Programs available for review on the ARS web site. Comments received continue to be distributed to the appropriate National Program teams for review and potential integration. Once the comment period has ended and subsequent revisions to the program statements are made, the IS will develop brochures aimed at ARS staff which will announce the new National Program Structure, address some of the frequently asked questions, and refer users to the NPS web site. Activities related to developing a long term strategy were held in abeyance until many of these short term activities were either well under way or completed.

At the first meeting of the full team, the charge of the action team was discussed and numerous background papers distributed. These papers included: a January 10, 1997 memo from G. Poulos and B. Kinzel, a January 17, 1997 memo from Ray Carruthers, and a February 4, 1997 memo from J. L. Hatfield concerning recommendations to and of the NPS Information Concept Working Group; October 1997 proposed models of ARS Internet data resources and a table of recommended action items with milestones from Pamela Mason; and information technology management in USDA and ARS handouts distributed by P. Andre at the September AC meeting. Other sources were identified and distributed to the team throughout the course of the project.

Vision and Mission of the Agricultural Research Service

Any initiative that requires a significant investment of resources should have a strong link to the vision and mission of the organization. Therefore, ARS' vision for "Leading America toward a better future through agricultural research and information" should be clearly visible through any strategic communication activity. This vision is supported through ARS' mission to "[c]onduct research to develop and transfer solutions to agricultural problems of high national priority and provides information access and dissemination to ensure high-quality, safe food, and other agricultural products, assess the nutritional needs of Americans, sustain a competitive agricultural economy, enhance the natural resource base and the environment, and provide economic opportunities for rural citizens, communities, and society as a whole." These roles of providing "information access and dissemination" and "transfer[ing] solutions to agricultural problems ..." are critical components of and the underlying premise beneath the overall Agricultural Research Service Communication Plan.

ARS Headquarters

Each of the background papers examined refers to a centralized repository as the authoritative source for information. This centralized source is seen as an internal working tool into and through which current and accurate program, budgetary, and administrative data are entered, viewed, manipulated, and extracted by authorized users. This repository also is seen as the source of information for externally focussed information products. ITM Action Team 4 who, in a more general sense, is charged with redesigning ARS program and resource management processes and systems also is charged with examining the nature and scope of this centralized repository.

The types of information products that can be gleaned from a centralized resource and that were identified as being of potential interest to ARS customers include: predefined descriptions of ARS activities in terms of specific commodities (such as corn, potatoes, dairy products, etc.) or technologies (such as genetic engineering, pesticide reduction, etc.) and matrices linking aspects of the National Programs to the actual research locations. Other information products of potential merit to ARS clients that should be managed centrally include: a dynamic contact list of experts; responses to frequently asked questions (FAQs) and a section on hot topics; and a centralized access mechanism for Internet-accessible ARS-generated research publications, data, and information. These products would be available via the ARS web site in a location determined by the ARS World Wide Web Board. They would serve as the organizational umbrella under which more detailed information could be obtained on specific research projects.

The data for the predefined commodity and technology descriptions and the National Program matrices could be extracted from the current Research Management Information System (RMIS) or redesigned centralized repository on an as needed or periodic basis. The frequency with which each of these products would be updated would be contingent upon the traditions of ARS administration and the messages the National Programs wish to communicate. The methods for updating the web site should be integrated as closely as possible into existing and/or revised processes and procedures for adding, deleting, or modifying program information.

The National Program Leaders (NPLs) would be the source of information for the contact list of experts, also known as the "Find the Expert" database. Personnel changes and shifts in program emphases would necessitate changes to the contact information. This update activity also could be done on an as needed basis or incorporated into a periodic review process.

The IS would be the logical source for FAQs and hot topics. Although the information reflected in these resources would be collected from many sources, some centralized level of oversight and coordination is needed to avoid duplication and to ensure the accuracy and currency of the information. A close relationship between IS and NPS would assure that relevant topics were being addressed. As these topics cool down, they should be archived and made accessible via a keyword and/or field delimited search facility.

Centralized Access Mechanism

At the September AC meeting, Lee Panella demonstrated a prototype web-based system created by the World Wide Web Agricultural Information Pilot Project Team (COWY Team) and designed to "facilitate subject access to work and accomplishments of current research units comprising the Colorado-Wyoming Research Council." At the heart of this system was a centralized database containing metadata records describing ARS-produced Internet-accessible resources. These resources ranged from computer models to research reports and publication lists. The demonstration also illustrated how an organized Web-based system could operate and the benefits of such a system to those attempting to locate information. The concept was well received by those in attendance with some voicing concern over the long-term support and funding of an ARS-wide implementation.

Many ARS locations already have made electronic resources available via their web sites. The key component that is presently lacking in the ARS web system is a centralized mechanism for accessing these resources. This mechanism could be developed either through the creation of a new automated administrative process or through the adaptation of an existing one.

Existing policies and procedures that could be adapted to develop a centralized point for accessing ARS-produced electronic resources are described in Directive 150.1 "Dissemination of Public Information by ARS," REE P&P 152-1 "Procedures for Publishing Manuscripts & Abstracts with Non-USDA Publishers (Outside Publishing)," and REE P&P 151.1 "Publishing (Print and Electronic)." These directives dictate the responsibilities and constraints of the various parties involved in disseminating ARS-related information as well as the approval process for this activity. They attempt to assure the quality of the information released to the public by ARS staff.

An integral component of the REE publishing policies and procedures is the creation of an ARS115 which tracks the publication approval process throughout the ARS organization. The RMIS system is used for the automated creation and modification of these forms and to facilitate the approval process. A by-product of these forms is the TEKTRAN database which includes

summaries “of recent research results” generated by ARS scientists and which is available both through RMIS (this version contains all the information submitted via the ARS115) and through the Internet-accessible TEKTRAN (this version contains a subset of the data submitted on the ARS 115) made available by the Technology Transfer Information Center (TTIC) at the National Agricultural Library. TEKTRAN is updated regularly to “includ[e] summaries of new articles that scientists have submitted for publication and remov[e] summaries after three years.” Mechanisms are also in place to protect potential patents by excluding related summaries from public view.

The Internet-accessible version of TEKTRAN has already proven to be a valuable resource for the public to obtain information on ARS-related research. Based on statistics gathered by NAL’s TTIC, TEKTRAN was accessed over 800,000 times in fiscal year 1997. The most popular topics, based on the ARS strategic planning code, included: pregnant and lactating women; pathogens and nematodes, nutrition; diseases; plant genetics and breeding; and naturally occurring toxic factors.

The relatively poor quality of Interpretive Summaries generated by ARS researchers for the ARS 115 and displayed in the TEKTRAN records was a source for concern by many. In order to increase the value of this database to ARS’ customers, serious consideration should be given to improve the quality of the records and the nature of the summaries. While addressing the data quality issue, the data elements should be reviewed to determine their compliance with existing and emerging metadata standards. These include the Dublin Core, those proposed in the REE Information System (REEIS), and the NBII (National Biological Information Infrastructure). If the scope of this database is expanded to include other electronic resources (such as databases) made available by ARS researchers via the Internet, then other standards such as GILS (Government Information Locator Service) also should be examined.

The benefits to adapting the present ARS 115 approval process and the resulting TEKTRAN database include: 1) internal ARS procedures for submitting data already exist and could be modified, if necessary; 2) the existing TEKTRAN data elements closely resemble a subset of those recommended by the COWY Team (i.e., *Authors* in TEKTRAN - *Investigators* in the recommended COWY database, *Interpretive Summary* in TEKTRAN - *Description* in COWY, *Keywords* in TEKTRAN - *Keywords* in COWY, etc.) and those in other existing and emerging metadata standards; and 3) review mechanisms already exist and could be strengthened to improve the meaningfulness of the data to ARS customers.

The creation of a centralized access mechanism can be achieved using any of three system models. These models are best described as centralized, hybrid, and decentralized. In a centralized system, both the metadata (one possible definition of the term “metadata” is: records describing a resource) and the electronic resources described by the metadata reside on and can be accessed from a centralized repository. Within a hybrid system, the metadata records reside on and can be accessed from a centralized repository but they link to and serve as access mechanisms for resources located on servers throughout the system. A decentralized system relies on search and

retrieval programs and scripts accessed through the centralized server which search metadata and/or data residing on disparate servers throughout the system.

The Centralized System

Using this model, all metadata and electronic resources would reside within a centralized server. The Staff Action system used by the Office of the Executive Secretariat for controlled correspondence is an example of this model. Images of the incoming correspondence, their descriptions, routing information, and the resulting responses are all accessible via a centralized system. Given sufficient storage and computing power, this model could be adapted for electronic publications via a re-engineered RMIS system. In addition to modifying the data elements within the ARS115 and improving the quality control process for the data, a carefully orchestrated and labor intensive effort would need to be undertaken to identify, obtain, load and centrally manage these electronic resources.

The Hybrid System

Within a hybrid system, metadata records reside on and can be accessed from a centralized repository and serve as access mechanisms for resources located on disparate systems. This approach could be adapted relatively quickly through modification of the present ARS115 requirements by asking the researcher to supply URLs (Uniform Resource Locator) for resources they have made accessible via the WWW. These URLs could be added to each field in the ARS115 for which additional information could be obtained from an ARS web site. Once the data are extracted from RMIS and loaded into the TTIC Internet-accessible TEKTRAN, the URLs could be activated so that the customer would be able to navigate to that resource without human intervention. The Area Offices would be responsible for assisting those locations without the necessary resources in making their publications Internet-accessible.

The Decentralized System

The Internet as a whole can be viewed as one large decentralized system from which search engines such as Alta Vista and OpenText glean information. Despite numerous initiatives, obtaining relevant high-quality information in such an uncontrolled environment has proven to be extraordinarily difficult. There exist, however, much more controlled decentralized systems which could serve as a model for ARS. One such system, the NASA Image eXchange (NIX) System (www.nix.gov), searches about 400,000 photos and data images from seven NASA Centers. The images and their descriptions reside on servers located at the geographically dispersed centers. Once a customer has initiated the search, it is sent simultaneously to each server. The results of the search are then returned from each server to the originating source, sorted and ranked, and made available to the searcher. NASA views the Image eXchange as the first step toward a comprehensive decentralized online imagery collection, and other collections will be added as they become available.

In order to implement an ARS-wide system using this model, each Area or other designated location would be responsible for developing and maintaining the system which would store and provide access to the publications produced within their jurisdiction. Each site could have its own

search and retrieval software as long as the data and/or metadata were stored in a format accessible via a centralized retrieval package. The centralized search mechanism would be prominently accessible via the ARS web site.

System Recommendation

In reviewing the present organizational structure and the state of technology within ARS, the implementation of a hybrid system appears to be the most cost effective and practical solution for increasing access to ARS information via the Internet. The modifications to the ARS115 should be an integral part of the RMIS re-design process and should be included as soon as possible. Some short term actions could be taken (such as adding URLs to data elements reflecting existing Internet accessible resources) while longer term redesign measures are being considered and implemented.

Summary of Action Item Recommendations and Activities

1. Post National Program Statements on ARS website for customer and stakeholder review. (Done)
2. Distribute letters to customer and stakeholders requesting review of statements. (Done)
3. Send announcement of the availability of the program statements via the ARS web site to Federal Register. (Done)
4. Create and disseminate brochures describing National Program Structure. (Pending)
5. Develop predefined commodity and technology descriptions for inclusion on ARS web site. (Recommendation)
6. Develop a dynamic contact list of experts for inclusion on ARS web site. (Done& Ongoing)
7. Develop a section for hot topics and responses to frequently asked questions (FAQs) for inclusion on ARS web site. Archive these as topics grow outdated. (Recommendation)
8. Develop a centralized access mechanism for Internet-accessible ARS publications and information based on the ARS115 process. Include modifications to this process in the RMIS redesign activities. Headquarters would be responsible for the centralized metadata repository and the ARS Area Offices would coordinate the posting of publications on the Internet for their respective locations. (Recommendation)

Appendix A
Schedule & Milestones

Task	Target Date	Actual Date
Charge & Background Material to Committee	Oct. 27, 1997	Oct. 27, 1997
Establish an NPS Web Presence	Dec. 19, 1997	Feb. 20, 1998
Distribute letters to stakeholders requesting review of statements	Dec. 19, 1997	Mar. 13, 1998
Send brief announcement of website to Federal Register	mid-Mar.1998	
Begin drafting long-range plan	Jan. 5, 1998	
Complete stakeholder evaluation period	Apr. 30, 1998	
Create and disseminate brochure and information packets	mid-June 1998	
National Program Teams complete review and finalize statements	end of May 1998	
Submit long-range plan	Apr. 30, 1998	

Time frame 6 months: 10/97 - 4/98
Proposed Budget: $5,000.00

Appendix B
Background Material